Weekly Report

***Period**: 05/04/2014 – 05/11/2014*

Projects

Research

As the number of data point increases, dissimilarity estimation dominates the computation consumptions in most multi-dimensional projection techniques. The classic MDS requires to compute the pairwise distance matrix as input. Other techniques like LSP, PLSP, and many other force based methods also need to compute the pairwise distance to build the KNN graph for projection. Therefore, we resort to the computationally efficient method called **Landmark MDS** (LMDS) to projection the high-dimensional fibers into a 2D visual space. The first step of the algorithm is to run classical MDS to embed a chosen subset of the data, referred to as the 'landmark points', in a low-dimensional space. Each remaining data point can be located within this space given knowledge of its distances to the landmark points. With this method, the distance computation complexity reduces to ***O****(n)*. In addition parallelism is another merit of this method, because each fiber to be projected only depends on the selected lardmark points.

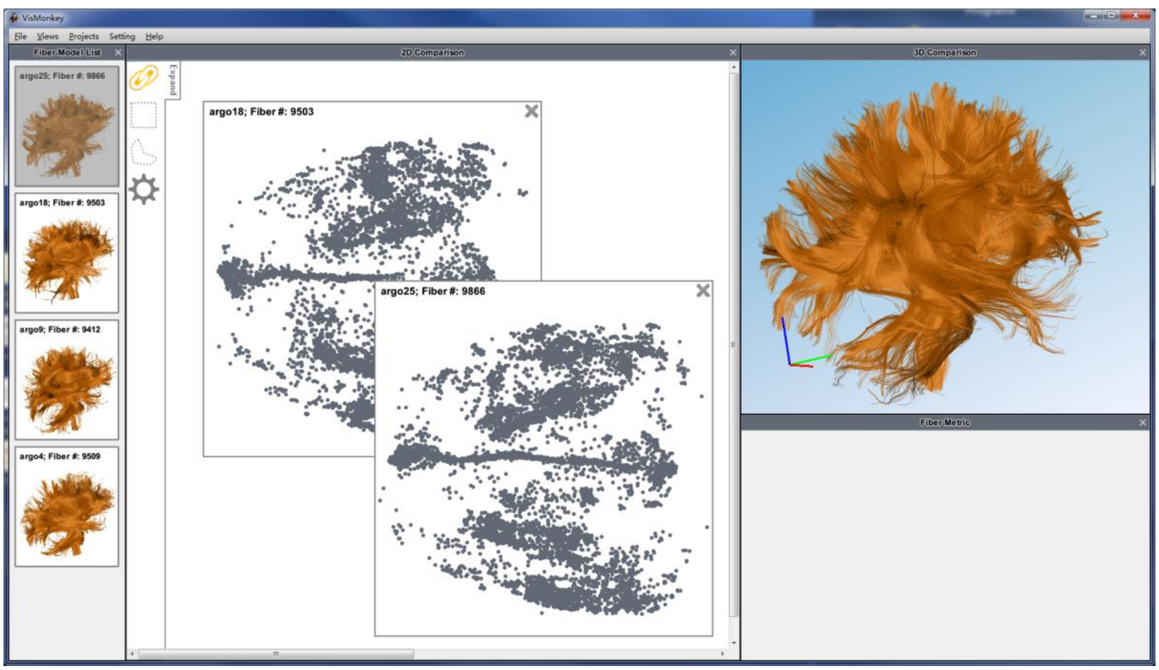

Figure 1 The main interface to explore differences of DTI fiber models

Figure 1 shows the main interface to study the differences among a set of fiber models. On the left side, a side-by-side glyph representation list view gives a high level overview for each fiber model. Users can drag a set of glyphs into the main view. The 2D

scatterplot representations can be further enhanced with KDE map to assist users study the differences (see Figure 2).

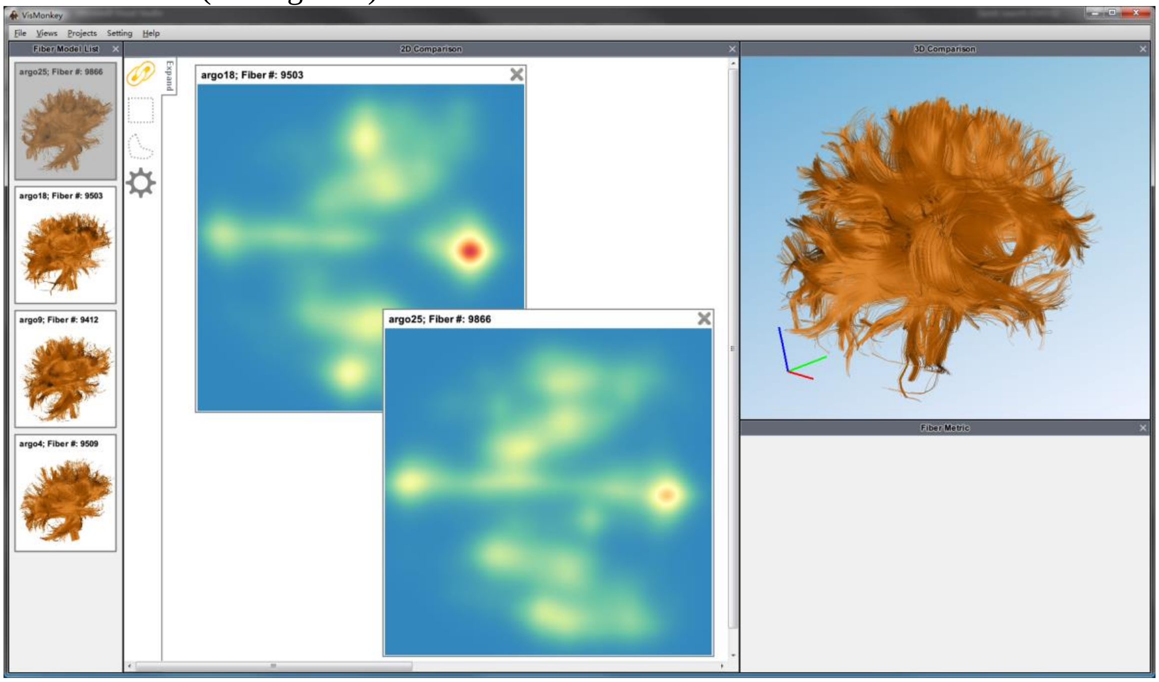

Figure 2 Using KDE to visualize the 2D projections of fiber models

The axes of the 2D projections does not have any intrinsic meanings. The information in the scatter-plot is very fine-grain and does not permit an intuitive understanding of its organization. To be able to convey the information on the screen to the user, we create an abstraction layer on top of the scatter-plot display.

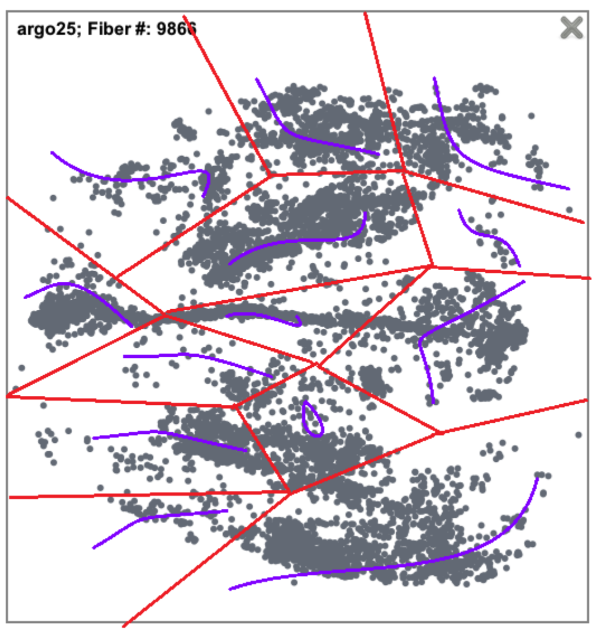

Figure 3 The mockup of the enhanced 2D embedding of fiber models

Specifically, we first randomly select a set of points in the embedding space for clustering. The cluster centroids are employed to tessellate the entire embedding space

into regions. The centroid fiber is selected as the representative fiber for each sub region and is rendered to the 2D plane (see Figure 3 as an example).

Work to be done in next week

- Implement the enhanced 2D projection for fiber models
- Implement interaction tools such as box selection, lasso selection

Reference: